



(19)

(11) Publication number: **10051755 A**

Generated Document.

**PATENT ABSTRACTS OF JAPAN**(21) Application number: **09047849**(51) Intl. Cl.: **H04N 7/15 G09G 5/36 H04M 3/56 H04N 5/232**(22) Application date: **03.03.97**(30) Priority: **30.05.96 JP 08136251**(43) Date of application  
publication: **20.02.98**(84) Designated contracting  
states:(71) Applicant: **FUJITSU LTD**(72) Inventor: **ENAMI TAKAFUMI**

(74) Representative:

**(54) SCREEN DISPLAY  
CONTROLLER FOR VIDEO  
CONFERENCE TERMINAL  
EQUIPMENT**

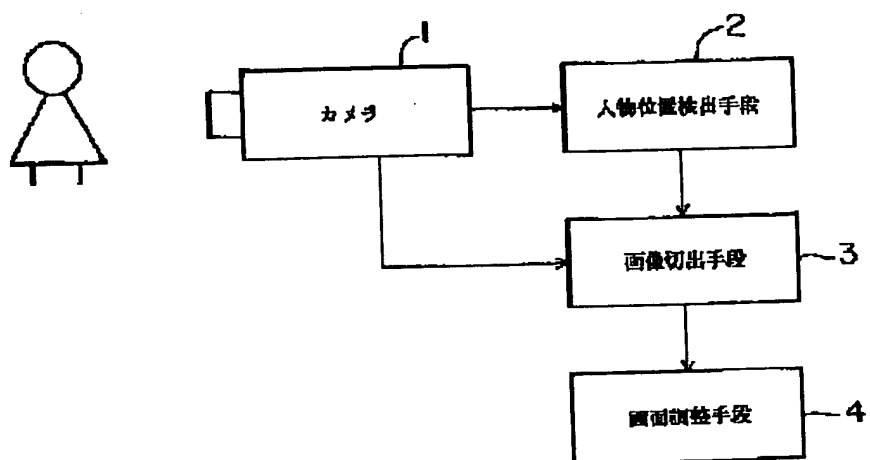
(57) Abstract:

**PROBLEM TO BE SOLVED:** To display an easy to see image automatically even when a person is moved by surely detecting a person in the case of displaying properly an image of person.

**SOLUTION:** A camera 1 sends a picked up broad angle image of an entire conference room including persons to a person position detection means 2 and an image segmentation means 3. The person position detection means 2 detects a position of a person based on a received image. The person image segmentation means 3 decides the range of image including the person to be displayed depending on the degree of movement of the person on an image detected by the position detection means 2 and segments an image range from the image of the camera 1. An image adjustment means 4 adjusts the image segmented by the image segmentation means 3 into a size of the screen of the image display device.

**BEST AVAILABLE COPY**

COPYRIGHT: (C)1998,JPO



(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平10-51755

(43) 公開日 平成10年(1998) 2月20日

(51) Int.Cl. <sup>9</sup>	識別記号	序内整理番号	F I	技術表示箇所
H 0 4 N 7/15			H 0 4 N 7/15	
G 0 9 G 5/36	5 2 0		G 0 9 G 5/36	5 2 0 P
				5 2 0 E
H 0 4 M 3/56			H 0 4 M 3/56	C
H 0 4 N 5/232			H 0 4 N 5/232	Z

審査請求 未請求 請求項の数27 O L (全 22 頁)

(21) 出願番号 特願平9-47849

(22) 出願日 平成9年(1997) 3月3日

(31) 優先権主張番号 特願平8-136251

(32) 優先日 平8(1996) 5月30日

(33) 優先権主張国 日本 (J P)

(71) 出願人 000005223

富士通株式会社

神奈川県川崎市中原区上小田中4丁目1番  
1号

(72) 発明者 枝並 隆文

神奈川県川崎市中原区上小田中4丁目1番  
1号 富士通株式会社内

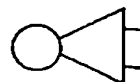
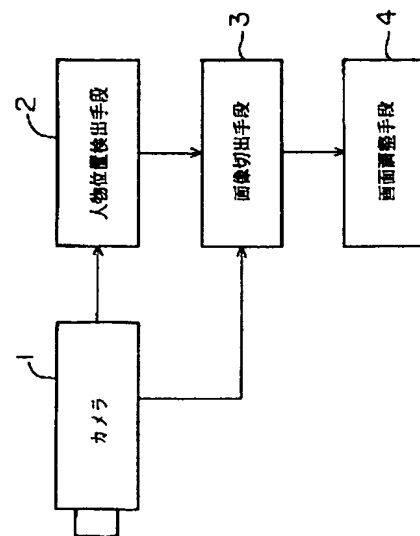
(74) 代理人 弁理士 服部 毅巖

(54) 【発明の名称】 テレビ会議端末の画面表示制御装置

(57) 【要約】

【課題】 人物画像の適切な表示を行うためのテレビ会議端末の画面表示制御装置に関し、人物検出を確実にを行い、かつ人物の移動があっても見やすい画像表示を自動的に行うことを課題とする。

【解決手段】 カメラ1が、人物を含めた会議室全体の広角撮影画像を人物位置検出手段2および画像切出手段3へ送る。人物位置検出手段2は、送られた画像を基にして人物位置を検出する。画像切出手段3が、人物位置検出手段2により検出された人物位置の画面上での移動度合いに応じて、人物を含めた表示されるべき画像範囲を決定し、カメラ1の画像から当該画像範囲を切り出す。画面調整手段4が、画像切出手段3により切り出された画像を画像表示装置の画面の大きさに調整する。



## 【特許請求の範囲】

【請求項1】 人物画像の適切な表示を行うためのテレビ会議端末の画面表示制御装置において、カメラから送られた広角撮影画像を基にして人物位置を検出する人物位置検出手段と、前記人物位置検出手段により検出された人物位置の画面上での移動度合いに応じて、前記人物を含めた表示されるべき画像範囲を決定し、前記カメラの画像から当該画像範囲を切り出す画像切出手段と、前記画像切出手段により切り出された画像を画像表示装置の画面の大きさに調整する画面調整手段と、を有することを特徴とするテレビ会議端末の画面表示制御装置。

【請求項2】 前記人物位置検出手段は、前記カメラから送られた画像を基にして人物の顔位置を検出する顔位置検出手段と、前記顔位置検出手段が検出した顔位置を基準にして前記人物の身体位置を検出する身体位置検出手段と、を含むことを特徴とする請求項1記載のテレビ会議端末の画面表示制御装置。

【請求項3】 前記顔位置検出手段は、前記カメラから送られた画像の中の色を、予め設定された顔色と比較して人物の顔位置である可能性がある部分を探索する顔色探索手段と、前記カメラから送られた画像の中で動く部分を探索する動き探索手段と、前記顔色探索手段および前記動き探索手段による両探索結果に基づき、人物の顔が存在する位置を判別する顔位置判別手段と、を含むことを特徴とする請求項2記載のテレビ会議端末の画面表示制御装置。

【請求項4】 前記顔位置検出手段は、前記カメラから送られた画像に対して、所定数の画素間隔のサンプリングを行うサンプリング手段を、更に含み、前記顔色探索手段および前記動き探索手段は、前記サンプリング手段によりサンプリングされた画像を基に各探索を行うことを特徴とする請求項3記載のテレビ会議端末の画面表示制御装置。

【請求項5】 前記動き探索手段は、所定時間間隔毎のフレームを基に動く部分を探索することを特徴とする請求項3記載のテレビ会議端末の画面表示制御装置。

【請求項6】 前記顔色探索手段は、前記カメラから送られた画像の中の色を、予め設定された複数の色標本と比較し、その類似度合いを求めて、人物の顔位置である可能性を判別することを特徴とする請求項3記載のテレビ会議端末の画面表示制御装置。

【請求項7】 前記顔色探索手段は、前記カメラから送られた画像全体の色の偏りを検出する色偏り検出手段と、

前記色偏り検出手段により検出された色の偏りを基に、前記カメラから送られた画像の色を補正する色補正手段と、を含むことを特徴とする請求項3記載のテレビ会議端末の画面表示制御装置。

【請求項8】 前記顔色探索手段は、人物の顔が存在する位置が正しく判別されたときに、人物の顔位置である可能性があるとされた部分の色を、次の顔色の標本として設定する顔色標本設定手段を、含むことを特徴とする請求項3記載のテレビ会議端末の画面表示制御装置。

【請求項9】 前記身体位置検出手段は、予め設定された人物の身体位置の特徴を保持する保持手段と、前記顔位置検出手段が検出した顔位置を基に、前記保持手段に保持された人物の身体位置の特徴を参照して前記人物の身体が存在する可能性のある領域を決定する領域決定手段と、前記領域決定手段で決定された領域における画像の動きを検出する動き検出手段と、前記動き検出手段によって動きが検出された画素をもって身体位置とする身体位置判定手段と、を含むことを特徴とする請求項2記載のテレビ会議端末の画面表示制御装置。

【請求項10】 前記画像切出手段は、前記人物位置検出手段により検出された人物位置の画面上での移動度合いを検出する移動検出手段と、前記移動検出手段で検出された移動度合いが大きいために、前記人物を含めた表示されるべき画像範囲を大きく設定し、移動度合いが小さいときに、前記人物を含めた表示されるべき画像範囲を小さく設定する画像範囲設定手段と、前記カメラの画像から、前記画像範囲設定手段により設定された画像範囲の画像を抽出する画像抽出手段と、を含むことを特徴とする請求項1記載のテレビ会議端末の画面表示制御装置。

【請求項11】 前記人物位置検出手段は、前記検出した人物位置に対して時間軸上でメディアンフィルタを作作用させるメディアンフィルタ作用手段を、含むことを特徴とする請求項1記載のテレビ会議端末の画面表示制御装置。

【請求項12】 人物画像の適切な表示を行うためのテレビ会議端末の画面表示制御装置において、カメラから送られた広角撮影画像を基にして複数の人物の顔位置および顔領域面積を検出する顔位置面積検出手段と、前記顔位置面積検出手段が検出した複数の人物の顔領域面積を比べ、最大面積の人物を選び、当該人物について検出された顔位置を基準にして当該人物の身体位置を検出する身体位置検出手段と、前記身体位置検出手段により検出された身体位置の画面

上での移動度合いに応じて、前記選ばれた人物を含めた表示されるべき画像範囲を決定し、前記カメラの画像から当該画像範囲を切り出す画像切出手段と、前記画像切出手段により切り出された画像を画像表示装置の画面の大きさに調整する画面調整手段と、を有することを特徴とするテレビ会議端末の画面表示制御装置。

【請求項13】 人物画像の適切な表示を行うためのテレビ会議端末の画面表示制御装置において、カメラから送られた広角撮影画像を基にして複数の人物の位置を検出する人物位置検出手段と、前記人物位置検出手段により検出された複数の人物位置の画面上での移動度合いに応じて、前記複数の人物を含めた表示されるべき画像範囲を決定し、前記カメラの画像から当該画像範囲を切り出す画像切出手段と、前記画像切出手段により切り出された画像を画像表示装置の画面の大きさに調整する画面調整手段と、を有することを特徴とするテレビ会議端末の画面表示制御装置。

【請求項14】 人物画像の適切な表示を行うためのテレビ会議端末の画面表示制御装置において、カメラから送られた広角撮影画像を基にして複数の人物の顔位置および顔領域面積を検出する顔位置面積検出手段と、前記顔位置面積検出手段が検出した複数の人物の顔領域面積を比べ、最大面積の人物を選び、当該人物についての検出された顔位置を基準にして当該人物の身体位置を検出する身体位置検出手段と、前記身体位置検出手段により検出された身体位置の画面上での移動度合いに応じて、前記選ばれた人物を含めた表示されるべき画像範囲を決定し、前記カメラの画像から当該画像範囲を切り出す第1の画像切出手段と、前記カメラから送られた画像を基にして複数の人物の位置を検出する人物位置検出手段と、前記人物位置検出手段により検出された複数の人物位置の画面上での移動度合いに応じて、前記複数の人物を含めた表示されるべき画像範囲を決定し、前記カメラの画像から当該画像範囲を切り出す第2の画像切出手段と、前記第1の画像切出手段および前記第2の画像切出手段が出力する画像のうちの一方の画像を選択して出力する選択出力手段と、前記選択出力手段から出力された画像を画像表示装置の画面の大きさに調整する画面調整手段と、を有することを特徴とするテレビ会議端末の画面表示制御装置。

【請求項15】 人物画像の適切な表示を行うためのテレビ会議端末の画面表示制御装置において、カメラから送られた広角撮影画像を基にして、フレーム間差分を算出するフレーム間差分算出手段と、前記広角撮影画像を基にして、フレーム内差分を算出す

るフレーム内差分算出手段と、

前記算出されたフレーム間差分及びフレーム内差分における重複部分を算出する重複部分算出手段と、前記算出された重複部分を基にして、前記広角撮影画像から人物画像を切り出す人物画像切出手段と、前記人物画像切出手段により切り出された人物画像を画像表示装置の画面の大きさに調整する画面調整手段と、を有することを特徴とするテレビ会議端末の画面表示制御装置。

【請求項16】 前記フレーム間差分のフレーム全体にわたる平均値を算出するフレーム間差分平均値算出手段を、さらに有し、前記フレーム間差分算出手段は、前記フレーム間差分平均値算出手段で算出された平均値を閾値として使用して、フレーム間差分を2値化するフレーム間差分2値化手段を含み、前記フレーム内差分算出手段は、前記フレーム間差分平均値算出手段で算出された平均値を閾値として使用して、フレーム内差分を2値化するフレーム内差分2値化手段を含むことを特徴とする請求項15記載のテレビ会議端末の画面表示制御装置。

【請求項17】 前記人物画像切出手段は、前記算出された重複部分を基にして、当該重複部分を構成する画素の中から、画面最上に位置する画素を検出するとともに、当該最上画素に連結する水平方向の画素を検出する最上行画素検出手段と、前記重複部分を構成する画素の中から、前記最上行画素検出手段で検出された連結画素群に連結する下方の画素を検出する下方画素検出手段と、前記連結画素群の直ぐ下の行において、前記下方画素検出手段で検出された画素を含み、前記連結画素群の幅以上の幅を有する行領域を設定する第1の行領域設定手段と、下方の行毎に、前記下方画素検出手段で検出された画素を含み、直ぐ上の行を構成する行領域の幅以上の幅を有する行領域を設定する第2の行領域設定手段と、前記重複部分を構成する画素の中から、前記最上行画素検出手段により検出された連結画素群と、前記第1及び第2の行領域設定手段により設定された各行領域とに含まれる各画素とを削除する削除手段と、を含むことを特徴とする請求項15記載のテレビ会議端末の画面表示制御装置。

【請求項18】 前記人物画像切出手段は、前記削除手段によって画素の削除が行われた後の重複部分を基にして、前記最上行画素検出手段、前記下方画素検出手段、前記第1の行領域設定手段、前記第2の行領域設定手段、および前記削除手段を繰り返し動作させ、前記最上行画素検出手段が画面最上に位置する画素を検出できなくなったときに前記繰り返し動作を停止させる動作制御手段をさらに含むことを特徴とする請求項17

記載のテレビ会議端末の画面表示制御装置。

【請求項19】 前記人物画像切出手段は、人物の形状に関する特徴を保持する特徴保持手段と、前記最上行画素検出手段により検出された連結画素群と、前記第1及び第2の行領域設定手段により設定された各行領域とから構成される検出形状を、前記特徴保持手段に保持されている特徴と比較して、前記検出形状が人物に相当する形状であるか否かを判定する判定手段と、をさらに含むことを特徴とする請求項17記載のテレビ会議端末の画面表示制御装置。

【請求項20】 前記特徴保持手段には、前記検出形状の上辺及び下辺の各長さ、並びに前記検出形状の高さにおける相互間の比に相当する実際の人物における比が保持されることを特徴とする請求項19記載のテレビ会議端末の画面表示制御装置。

【請求項21】 前記人物画像切出手段は、前記最上行画素検出手段により検出された連結画素群と、前記第1及び第2の行領域設定手段により設定された各行領域とから構成される検出形状を基に、人物の顔位置を検出する顔位置検出手段をさらに含むことを特徴とする請求項17記載のテレビ会議端末の画面表示制御装置。

【請求項22】 前記人物画像切出手段は、人物の顔位置の色パターンを保持する色パターン保持手段と、前記顔位置検出手段により検出された顔位置における色パターンを、前記色パターン保持手段に保持された色パターンと比較する色パターン比較手段と、前記色パターン比較手段による比較結果に基づき、人物の検出を確認する人物検出確認手段と、をさらに含むことを特徴とする請求項21記載のテレビ会議端末の画面表示制御装置。

【請求項23】 前記色パターン保持手段は、常時、信頼性の高い色パターンを保持するようにすることを特徴とする請求項22記載のテレビ会議端末の画面表示制御装置。

【請求項24】 前記人物画像切出手段は、画像表示装置に表示すべき人物を特定するための複数の方法モードのうちの1つを、外部から指定する指定手段を、さらに有し、前記人物画像切出手段は、前記指定手段により指定された方法モードに従い、前記広角撮影画像から人物画像を切り出すことを特徴とする請求項17記載のテレビ会議端末の画面表示制御装置。

【請求項25】 前記複数の方法モードは、全員切り出しモードと、個人切り出しモードとからなることを特徴とする請求項24記載のテレビ会議端末の画面表示制御装置。

【請求項26】 前記人物画像切出手段は、

個人切り出しモードにおいて使用され、各人物の動き量に基づいて特定の人物の画像を自動的に切り出す第1の人物画像切出手段をさらに有することを特徴とする請求項25記載のテレビ会議端末の画面表示制御装置。

【請求項27】 前記人物画像切出手段は、個人切り出しモードにおいて使用され、各人物の画面上での面積に基づいて特定の人物の画像を自動的に切り出す第2の人物画像切出手段をさらに有することを特徴とする請求項25記載のテレビ会議端末の画面表示制御装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、テレビ会議端末の画面表示制御装置に関し、特に、人物画像の適切な表示を行うためのテレビ会議端末の画面表示制御装置に関する。

【0002】テレビ会議端末においては、相手端末で再生される画像に自端末側の人物が正常に写っているかどうかを確認することが必要である。これを行わないとせっかくテレビ会議を行っているにもかかわらず、相手に送られた画像に人物が含まれていなかったり、人物の一部しか含まれていないことがしばしば発生し、円滑なテレビ会議の進行を阻害する。本発明は、そうした人物像の確実な送信に関する技術を提供する。

【0003】

【従来の技術】従来の最も一般的な装置では、相手端末へ送信している画像を、自端末側の画像表示装置の画面の一部にもモニタ画像として表示するようにしている。この装置では、送信側において、モニタ画面に写っている人物自身がその画面を見ながら定期的に自分の姿が正常に送られているか、すなわち、画面の中に正常に入っているかどうかを常時確認し、画面から自分の画像がはみ出しているような場合には、手動でカメラの向きを修正するなどの行為が必要となる。そのため、自分の姿が正常に送られているかどうかを常に気を使っていなければならない、またカメラの位置を修正したりすることによりテレビ会議の円滑な進行を阻害するという問題を抱えている。

【0004】また、カメラ自身にズーム機構および首振り機構を持たせたものを使用し、首振りやズームの操作を相手端末側や自端末側で行うようにした装置もある。この装置では、こうした操作を行う操作者が必要となる。特に操作者が会議参加者である場合には操作に意識が集中してしまい、円滑な会議の進行の妨げとなる。

【0005】こうした上記の装置ではいずれも、送信側の人物像が受信端末に正しく表示されるようにするために、注意を集中しなければならない人や操作を行わなければならない人が必要である。こうした人の介在を無用とすることを可能にした装置として、従来、送信側において、広角で解像度の高い固定カメラにより人物を含め

て会議室全体を撮影し、その画像を基にして人物の存在する領域を検出し、人物を中心に据えた一定の大きさの画像を切り取り、受信側の画像表示装置の画面に合わせた大きさに拡大して受信側へ送るようにしたものがあ  
る。この装置では、人物が移動しても、その移動が、広角の固定カメラで撮影されている範囲内であれば、受信側の表示装置の画面中央に常時その人物像を表示することが可能である。

#### 【0006】

【発明が解決しようとする課題】しかし、上記の人物像を受信側の画像表示装置の画面中央に自動的に常時表示する装置では、後述のように人物像の検出に問題があるとともに、人物像の検出ができたとしても、人物像の大きさを一定にし、それを中心に据えた画像を切り出すため、人物の移動に伴い背景のみが変化する。こうした画像では、背景を含めた人物の全体的な動きを把握することが困難になると共に、背景に関してはカメラ移動を過度に行なったときと同じような感じの画像となり、違和感のある見づらい画面表示になってしまう。

【0007】なお従来の人物検出の方法としては、はじめに人物のいない状態の背景画像を記憶しておき、その画像とテレビ会議中の人物を含む画像との差分を求めることにより人物位置を求める方法が知られている。しかし、この方法は、背景に時間的な変化が生じない場合にしか適用できず、また、この方法ではカメラの向きや位置を変えることもできないことになる。

【0008】また、従来の人物検出方法として、画像フレーム間の差分（動き量）を求めることで人物位置を抽出する方法もあるが、フレーム間の細かなノイズや背景の動きによるノイズなどが混在してしまい、十分安定した検出精度を得ることは難しい。

【0009】本発明はこのような点に鑑みてなされたものであり、人物検出を確実にを行い、かつ人物の移動があっても見やすい画像表示を自動的に行うことを可能にしたテレビ会議端末の画面表示制御装置を提供することを目的とする。

#### 【0010】

【課題を解決するための手段】本発明では上記目的を達成するために、図1に示すように、カメラ1から送られた広角撮影画像を基にして人物位置を検出する人物位置検出手段2と、人物位置検出手段2により検出された人物位置の画面上での移動度合いに応じて、人物を含めた表示されるべき画像範囲を決定し、カメラ1の画像から当該画像範囲を切り出す画像切出手段3と、画像切出手段3により切り出された画像を画像表示装置の画面の大きさに調整する画面調整手段4とを有することを特徴とするテレビ会議端末の画面表示制御装置が提供される。

【0011】以上のような構成において、カメラ1が、人物を含めた会議室全体の広角撮影画像を人物位置検出手段2および画像切出手段3へ送る。カメラ1は、固定

された広角ビデオカメラである。人物位置検出手段2は、送られた画像を基にして人物位置を検出する。具体的には、カメラ1からの画像を基にして、まず人物の顔位置を検出し、検出した顔位置を基準にして人物の身体位置を検出する。これらの顔位置や身体位置の検出には、顔色や動きベクトルが利用される。

【0012】つぎに、画像切出手段3が、人物位置検出手段2により検出された人物位置の画面上での移動度合いに応じて、人物を含めた表示されるべき画像範囲を決定し、カメラ1の画像から当該画像範囲を切り出す。すなわち、人物位置の画面上での移動度合いを検出し、この移動度合いが大きいときには、人物を含めた表示されるべき画像範囲を大きく設定し、一方、移動度合いが小さいときには、人物を含めた表示されるべき画像範囲を小さく設定する。そして、カメラ1の画像から、この設定された画像範囲の画像を抽出する。

【0013】最後に、画面調整手段4が、画像切出手段3により切り出された画像を画像表示装置の画面の大きさに調整する。このようにして、人物抽出を確実にを行い、かつ人物の移動があっても、自動的にズームアウトする方法により見やすい画像表示を可能としている。

#### 【0014】

【発明の実施の形態】以下、本発明の実施の形態を、図面を参照して説明する。まず、本発明に係るテレビ会議端末の画面表示制御装置の第1の実施の形態の原理構成を、図1を参照して説明する。本実施の形態は、カメラ1から送られた広角撮影画像を基にして人物位置を検出する人物位置検出手段2と、人物位置検出手段2により検出された人物位置の画面上での移動度合いに応じて、人物を含めた表示されるべき画像範囲を決定し、カメラ1の画像から当該画像範囲を切り出す画像切出手段3と、画像切出手段3により切り出された画像を画像表示装置の画面の大きさに調整する画面調整手段4とから構成される。

【0015】以上のような構成において、カメラ1が、人物を含めた会議室全体の広角撮影画像を人物位置検出手段2および画像切出手段3へ送る。カメラ1は、固定された広角ビデオカメラである。人物位置検出手段2は、送られた画像を基にして人物位置を検出する。具体的には、カメラ1からの画像を基にして、まず人物の顔位置を検出し、検出した顔位置を基準にして人物の身体位置を検出する。これらの顔位置や身体位置の検出には、顔色や動きベクトルが利用される。

【0016】つぎに、画像切出手段3が、人物位置検出手段2により検出された人物位置の画面上での移動度合いに応じて、人物を含めた表示されるべき画像範囲を決定し、カメラ1の画像から当該画像範囲を切り出す。すなわち、人物位置の画面上での移動度合いを検出し、この移動度合いが大きいときには、人物を含めた表示されるべき画像範囲を大きく設定し、一方、移動度合いが小

さいときには、人物を含めた表示されるべき画像範囲を小さく設定する。そして、カメラ1の画像から、この設定された画像範囲の画像を抽出する。

【0017】最後に、画面調整手段4が、画像切出手段3により切り出された画像を画像表示装置の画面の大きさに調整する。このようにして、人物抽出を確実に、かつ人物の移動があっても、自動的にズームアウトする方法により見やすい画像表示を可能としている。

【0018】図2は第1の実施の形態の詳しい構成を示すブロック図である。なお、図1に示す構成と図2に示す構成との対応関係については、図2に基づく説明を行った後に記述する。

【0019】図中、カメラ11が、会議に参加している人物を含めた会議室全体を撮影し、その画像をフレームメモリ12へ送り、格納する。カメラ11は、広範囲に撮影することができる広角のテレビカメラであるとともに、高解像度のテレビカメラである。ズーム機構や首振り機構を備えてはいず、カメラ11の向きや位置も固定されている。

【0020】フレームメモリ12に蓄えられた画像情報に対して、所定数の画素間隔のサンプリングが行われ、間引きフレームメモリ13に蓄えられる。顔位置検出回路14は、間引きフレームメモリ13に蓄えられた画像情報を処理することにより人物の顔位置や大きさを検出する。その顔位置や大きさを基にして、人物位置検出回路15が、人物の身体の位置の検出を行う。こうして検出された人物位置を基にして、切り出し位置決定回路16が、人物位置が移動している場合には、その人物像を含めた画像の切り出されるべき範囲を広く設定し（後述のようにズームアウトしたような画像になる）、人物位置が移動していない場合にはその画像の切り出されるべき範囲を狭く設定する（後述のようにズームインしたような画像になる）。

【0021】画像切り出し回路17は、設定された範囲

$$Q_{i,j}(n) = P_{sx, sy}$$

【0026】ただし、nはフレーム番号、sxはx方向間引き間隔、syはy方向間引き間隔である。なお、原画像21の画像データはカラー情報であるので、PおよびQはRGBデータまたはYC表現による3次元ベクトル値である。

【0027】こうしたサンプル点配列データ22のサンプル点 $Q_{i,j}(n)$ を用いて、顔位置検出回路14は、

$$E_{ci,j}(n) = \frac{\max_{k=0 \dots Nd} W(k) |Q_{i,j}(n) - D_c(k)|}{\dots} \quad (2)$$

【0029】ここで、 $D_c(k)$ は、前もって設定された人物の顔色（肌色）に関する標準値であり、Ndは標

準値の個数、 $| |$ は絶対値、 $W(k)$ は標準値毎の分散を基本とする重み係数である。

【0022】なお、図1に示すカメラ1は図2のカメラ11に対応し、図1に示す人物位置検出手段2は図2の顔位置検出回路14および人物位置検出回路15に対応し、図1に示す画像切出手段3は図2の切り出し位置決定回路16および画像切り出し回路17に対応し、図1に示す画面調整手段4は図2の画像拡大縮小回路18に対応する。

【0023】図3は顔位置検出回路14の動作を説明する図である。顔位置検出回路14は、色識別を基本とする顔位置検出処理（処理23、データ24）とフレーム間微分を基本とする顔位置検出処理（処理25、データ26）を行い、それらの結果を用いて人物の顔位置を特定する（データ27、データ28）。それらの処理を高速に行うために、フレームメモリ12に格納された原画像21の全画素に対してではなく、所定数の画素間隔のサンプル点（サンプル点配列データ22）に対してこうした処理を行うようにする。そのために、顔位置検出回路14の処理に先立って、データのサンプリングが行われる。

【0024】すなわち、原画像21の左上を起点とする画像データ配列を $P_{x,y}$ とすると、サンプル点配列データ22のサンプル点 $Q_{i,j}(n)$ を、下記式（1）に基づき算出する。

【0025】

【数1】

$$Q_{i,j}(n) = P_{sx, sy} \quad (1)$$

色識別を基本とする顔位置検出処理を行う。すなわち、下記式（2）に基づき、サンプル点 $Q_{i,j}(n)$ と、人物顔色の標準辞書の中に登録されている色との距離を、色評価値 $E_{ci,j}$ として算出する（処理23）。

【0028】

【数2】

準値の個数、 $| |$ は絶対値、 $W(k)$ は標準値毎の分散を基本とする重み係数である。



【0030】こうして得られた色評価値 $E_{ci,j}$ が、図3のデータ24である。ここでは、色評価値 $E_{ci,j}$ が大きく、顔色に近い点を大きな黒丸で図示している。なお、カメラ11の色入力値の偏りや、照明などの影響を考慮

$$Q'_{ci,j} = Q_{i,j} - S_c \quad \text{----- (3)}$$

【0032】ただし、 $S_c$  は画面内の平均色を示すオフセット量であり、下記式(4)で表せる。

$$S_c = \sum_{i=0}^{N_x-1} \sum_{j=0}^{N_y-1} Q_{i,j} / (N_x \cdot N_y) \quad \text{----- (4)}$$

【0034】また、サンプル点配列データ22のサンプル点 $Q_{i,j}$  ( $n$ ) を用いて、顔位置検出回路14は、フレーム間微分を基本とする顔位置検出処理を行う。すなわち、下記式(5)に基づき、フレーム間差分である動

$$E_{mi,j}(n) = |Q_{i,j}(n) - Q_{i,j}(n-1)| + a_1 \cdot E_{mi,j}(n-1) \quad \text{----- (5)}$$

【0036】ただし、 $N_p$  は差分を取るための次数、 $a_k$  は予測のための係数である。例えば  $N_p = 1$ 、 $a_1 = 1$  とすると、単純なフレーム間差分を計算することになる。 $N_d$ 、 $a_k$  は、実際には画像のフレーム間の時間により指定されることが望ましい。

【0037】こうして得られた動き評価値 $E_{mi,j}$ が、図3のデータ26である。ここでは、動き評価値 $E_{mi,j}$ が

$$E_{i,j} = A_c E_{ci,j} + A_m E_{mi,j} \quad \text{----- (6)}$$

【0040】ここで、 $A_c$ 、 $A_m$  はそれぞれ、色評価値 $E_c$  および動き評価値 $E_m$  の線形結合用の重み係数である。なお、動き評価値 $E_m$  は基本的に微分操作を含んでいるため、顔の周辺部分のみで大きくなり、顔の中心部では小さくなる傾向がある。こうしたことを避けるため、動き評価値 $E_m$  を隣接周辺部分に伝搬し、安定した顔の

$$E_{i,j} = A_c E_{ci,j} + \sum_{x=-1}^1 \sum_{y=-1}^1 A_m(x,y) E_{mi+x,j+y} \quad \text{----- (7)}$$

【0042】ここで、 $A_m(i,j)$  は動き評価値 $E_m$  を加算するための加算重み係数値であり、ガウス分布またはピラミッド型の重み係数として設定される。このようにして各サンプル点毎に計算された総合評価値 $E_{i,j}$  を用いて、顔位置検出回路14は顔位置の検出を行う。すなわ

$$E_{i,j} > TH_r \quad \text{----- (8)}$$

【0044】ここで、 $TH_r$  は顔位置の検出のための評価基準値であり、前もって設定されるものである。このようにして得られた顔候補点の座標をパラメータとするクラスタリング処理またはヒストグラム(投影)によるグルーピング処理を行うことにより、顔候補点の座標を数個のグループに分割する。このグループ毎の点群に対する外接四角形を計算することにより顔位置および顔の大きさを得ることができる(データ28)。

【0045】以上の処理により一度顔位置を検出した

して、上記の色評価値の代わりに、下記式(3)で表されるような $Q'_{i,j}$ を使用するようにしてもよい。

【0031】

【数3】

【0033】

【数4】

きに対する評価値 $E_{mi,j}$ を算出する(処理25)。

【0035】

【数5】

大きく、したがって人物像である可能性の高い点を大きな黒丸で図示している。

【0038】以上の色評価値 $E_c$  および動き評価値 $E_m$  を基にして、総合的な顔位置の評価値 $E_{i,j}$  を下記式(6)により定義する。

【0039】

【数6】

検出を実現することが望ましい。そのためには、下記式(7)に基づき「ぼかし」処理を施した総合評価値 $E_{i,j}$  を得るようにする。

【0041】

【数7】

ち、全サンプル点について、下記式(8)を満たす点を顔候補点として登録する(データ27)。

【0043】

【数8】

後、その時に検出された顔色の値をその後の顔位置検出処理の際の標本色に設定する。これにより、顔位置検出処理の精度の向上を図ることが可能である。

【0046】図4は人物位置検出回路15の動作を示す図である。人物位置検出回路15には、顔位置検出回路14から、図3のデータ26およびデータ28が送られる。これらを図4では、データ33およびデータ31とそれぞれ表すことにする。また人物位置検出回路15には、人物の顔位置に対する身体位置の形状を知識データ

ベース32として予め記憶しておく。これらのデータ26、データ28、知識データベース32を基にして、人物位置検出回路15は人物位置の検出を行う(処理34)。

【0047】すなわち、まず、知識データベース32として、顔位置の中心を基準とする身体位置の相対座標群を座標列により指定する。または、そうした相対座標群を、判定関数と、それに付随する評価値計算のための係数または係数式とにより定義する。このとき、身体位置

$$y < a_0 x$$

$$y < a_1 x$$

【0050】人物位置検出回路15は、こうした範囲を人物の身体が存在する可能性のある場所と見做して、データ33を基に、その範囲における動き評価値 $E_{mi,j}$ が大きい場所を探索する。

(b) 人物の身体形状を四角形で近似する場合には、人

$$a_0 < x < a_1 \quad (a_0 < a_1) \quad \text{----- (10a)}$$

$$y > b_0 \quad (b_0 < 0) \quad \text{----- (10b)}$$

【0052】人物位置検出回路15は、こうした範囲を人物の身体が存在する可能性のある場所と見做して、データ33を基に、その範囲における動き評価値 $E_{mi,j}$ が大きい場所を探索する。

(c) 人物身体形状を重み付き三角形領域で定義する場合

$$Cb = A \cdot a_0 + B \quad \text{----- (11)}$$

【0054】ただし、A、Bは定数であり、 $a_0$ は下記式(12)により算出する。

$$a_0 = |\tan^{-1}((y - y_f) / (x - x_f))| \quad \text{----- (12)}$$

【0056】ここで、 $x_f$ 、 $y_f$ は顔の中心座標値である。そして、こうして得られた重み $Cb$ を用いて、下記式(13)に基づき動き評価値 $E'_{mi,j}$ を算出する。

$$E'_{mi,j} = Cb \cdot E_{mi,j} \quad \text{----- (13)}$$

【0058】この動き評価値 $E'_{mi,j}$ のうちで、下記式(14)を満たす値を以て、人物の身体が存在する位置とする。

【0059】

【数14】

$$E'_{mi,j} > TH_b \quad \text{----- (14)}$$

【0060】ただし、 $TH_b$ は身体位置の検出のための評価基準値であり、前もって設定されるものである。動

$$E_{mi,j} > TH_b \quad (i, j \in (a) \text{ or } (b)) \quad \text{----- (15)}$$

【0063】すなわち、上記の場合(a)または場合(b)の場合、人物の身体が存在する可能性のある場所の中で、動き評価値 $E_{mi,j}$ が規定値 $TH_b$ を越える部分を身体位置とする。実際は、図4のデータ35に示すように、上記式(14)または(15)を満たす値からなる人物身体位置および顔位置を含む外接四角形をもって

は顔の大きさをパラメータとして決定される。

【0048】具体的には、顔位置を基準とする身体位置の相対座標群は次の様に表現される。

(a) 人物の身体形状を三角形で近似する場合には、人物の顔位置を座標中心(原点)と考えたときに、人物の身体形状を下記式(9a)、(9b)を満たす範囲として表すことができる。

【0049】

【数9】

$$\text{----- (9a)}$$

$$\text{----- (9b)}$$

物の顔位置を座標中心(原点)と考えたときに、人物の身体形状を下記式(10a)、(10b)を満たす範囲として表すことができる。

【0051】

【数10】

$$\text{----- (10a)}$$

$$\text{----- (10b)}$$

合には、まず、下記式(11)により身体位置に対応する重み $Cb$ を算出する。

【0053】

【数11】

【0055】

【数12】

$$\text{----- (12)}$$

【0057】

【数13】

$$\text{----- (13)}$$

き評価値 $E'_{mi,j}$ を実際の身体位置の判定パラメータとすることにより、安定に身体位置の検出を実現することが可能となる。

【0061】上記の場合(a)または場合(b)の場合には、下記式(15)を満たす値を以て、人物の身体が存在する位置とする。

【0062】

【数15】

$$\text{----- (15)}$$

人物位置とする。人物位置検出回路15は、顔中心位置、身体位置、頭部大きさ、人物外接四角形位置に関する情報を切り出し位置決定回路16へ送る。

【0064】なお、人物位置の検出にあたり、複数の人物が存在すると判断され得るサンプル点グループが存在する場合には、複数人物を包含する領域を切り出し領域

とるようにしてもよい。また、顔の面積を比べ、面積最大の人物のみを画像切り出し対象としてもよい。さらに、選択スイッチにより、こうした両方の方式を備え、それらのうちのいずれか一方の方式を選択できるようにしてもよい。

【0065】つぎに、切り出し位置決定回路16が人物画像切り出し位置の決定を行う。その決定の手順を、図5を参照して説明する。図5は切り出し位置決定回路16による人物画像切り出し位置の決定手順を示すフローチャートである。以下、図中のステップに沿って説明する。

【0066】〔S1〕身体位置パラメータの初期化を行う。その後、各フレーム毎または数フレーム毎に以下の処理を行う。各フレーム毎に処理を行うこととすると、処理が頻繁になるため細かな身体の動きやノイズなどの影響で切り出し位置が頻繁に動いてしまい、人物画像として非常に不安定で見にくいものとなる場合がある。そういった場合に数フレーム毎の処理とすることで画像を見やすいものにすることが可能である。

【0067】〔S2〕現在の人物位置情報を人物位置検

$$S_L' = \left( \text{MID}_{i=0, N_a, N_b} X_L(n-i) \right) - \alpha_L \quad \text{----- (16a)}$$

$$S_R' = \left( \text{MID}_{i=0, N_a, N_b} X_R(n-i) \right) + \alpha_R \quad \text{----- (16b)}$$

$$S_U' = \left( \text{MID}_{i=0, N_a, N_b} Y_U(n-i) \right) - \alpha_U \quad \text{----- (16c)}$$

$$S_D' = \left( \text{MID}_{i=0, N_a, N_b} Y_D(n-i) \right) + \alpha_D \quad \text{----- (16d)}$$

【0071】ここで、MIDはメディアンフィルタリング処理を示し、 $N_a$ 、 $N_b$ はそれぞれメディアンフィルタ個数と中央値個数を示す。この式により計算された $S_L'$ 、 $S_R'$ 、 $S_U'$ 、 $S_D'$ を元にして、下記式(17a)、(17b)、(17c)、(17d)に基づ

$$S_L = \text{MIN}_{i=0, N_c} (S_L') \quad \text{----- (17a)}$$

$$S_R = \text{MAX}_{i=0, N_c} (S_R') \quad \text{----- (17b)}$$

$$S_U = \text{MIN}_{i=0, N_c} (S_U') \quad \text{----- (17c)}$$

$$S_D = \text{MAX}_{i=0, N_c} (S_D') \quad \text{----- (17d)}$$

【0073】ここで、 $N_c$ は $S_L$ 、 $S_R$ 、 $S_U$ 、 $S_D$ の最大、最小を求める際のサンプル時間間隔である。

〔S3〕人物位置検出回路15から送られた顔中心位置、身体位置等を用いて、所定時間に人物が画面内で移動する量を算出する。

【0074】〔S4〕その算出された移動量を所定値と比較し、移動量が所定値よりも大きいときにはステップS5へ進み、移動量が所定値よりも小さいときにはステップS7へ進む。

出回路15から読み込み、それに基づいて画像切り出し領域の決定を行う。人物位置検出回路15からは顔中心位置、身体位置、頭部大きさ、人物外接四角形位置などの情報がもたらされるが、ここでは人物外接四角形位置を用いて画像切り出し領域の決定を行う。

【0068】すなわち、図6に示すように、人物外接四角形41の位置は $X_L$ （左側）、 $X_R$ （右側）、 $Y_U$ （上側）、 $Y_D$ （下側）によって表現される。ところで、頻繁に画像切り出し領域の変更を行わないようにするためには、画像切り出し領域を、人物外接四角形領域が安定に包含されるように設定することが望ましい。そのために、画像切り出し領域を $S_L$ （左側）、 $S_R$ （右側）、 $S_U$ （上側）、 $S_D$ （下側）とした場合に、これらを次のように決定するようにする。

【0069】まず、 $X_L$ 、 $X_R$ 、 $Y_U$ 、 $Y_D$ に対して、下記式(16a)、(16b)、(16c)、(16d)に基づきメディアンフィルタリング処理を行い、 $S_L'$ 、 $S_R'$ 、 $S_U'$ 、 $S_D'$ を算出する。

【0070】

【数16】

き、画像切り出し領域 $S_L$ 、 $S_R$ 、 $S_U$ 、 $S_D$ を決定する。

【0072】

【数17】

【0075】〔S5〕ステップS4の判別が行われた今回よりも前の移動量を所定値と比較し、移動量が所定値よりも大きいときにはステップS2へ戻り、移動量が所定値よりも小さいときにはステップS6へ進む。

【0076】〔S6〕過去に人物の移動量が多くなかったため、画像切り出し領域が狭く設定されているはずである。そうした状態において、今回、移動量が多い場合には、その人物が切り出された画面の外に出てしまう可能性があるため、画像切り出し領域を、移動量に応じた

量だけ広く設定する。すなわち、移動する人物を追尾するのではなく、結果的にズームアウトしたような画像表示がなされるように画像切り出し領域を設定する。

【0077】〔S7〕ステップS4の判別が行われた今回よりも前の移動量を所定値と比較し、移動量が所定値よりも小さいときにはステップS2へ戻り、移動量が所定値よりも大きいときにはステップS8へ進む。

【0078】〔S8〕過去に人物の移動量が多かったので、画像切り出し領域が広く設定され、人物像が小さく画面表示されているはずである。そうした状態において、今回、移動量が少ない場合には、その人物があまり動かなくなったと思われるので、画像切り出し領域を移動量に応じた量だけ狭く設定する。すなわち、結果的に人物に対してズームインしたような画像表示がなされるように画像切り出し領域を設定する。

【0079】〔S9〕ステップS6またはステップS8の設定に従い、最終切り出し領域を指定する。

〔S10〕最終切り出し領域と現在の切り出し領域との間に中間的な切り出し領域を指定する。この中間切り出し領域は、画像圧縮回路20での画像符号化において符号情報量が極端に増加しないようにするために設定するものであり、その点を考慮して段階的に変化する複数の中間切り出し領域を設定するようにする。

【0080】〔S11〕中間切り出し領域に基づく画像切りだしを行なった結果、最終切り出し領域に達したか否かを判別し、最終切り出し領域に達していればステップS2へ戻り、達していなければステップS10へ戻る。

【0081】図7は、送信側から送られた符号化画像を受信側で復号化して画像表示装置に表示した場合の画面の例を示す図である。すなわち、カメラ11により原画像42が撮影されたとする。通信開始時においては、受信側で画面43のようにほぼ全体画像を表示するようにする。次に、原画像42において人物に動きがなく安定している場合には、画像切り出し領域を狭くし、画面44のようなズームイン画像にする。その後、さらに人物が安定している場合には、画像切り出し領域をさらに狭くし、画面45のようなズームイン画像にする。なお、画面43から画面44へ、さらに画面45への変化は画面を見る人にとって違和感のない滑らかさと速度とで行う。

【0082】ここで、画面45のズームイン状態において人物が移動した場合には、画像切り出し領域を大きくし、移動した人物を包含する位置にまで大きくする。その結果、画面44さらには画面43へ移行する。

【0083】以上のように、人物の顔位置を顔の色情報と顔の動き情報とにより検出し、その顔位置を基にして身体位置の検出を行っている。身体位置の検出には、顔位置を基準として探索範囲を決定し、その探索範囲における動き検出によって身体位置を検出している。このた

め、背景中に移動する物があっても人物像の抽出を確実に行うことができる。また、人物が移動した場合に、人物画像の追跡を行うのではなく、ズームアウトと等価な画像切り出しを行うようにする。これにより、非常に見やすいテレビ会議画像の表示を実現している。

【0084】次に、第2の実施の形態を説明する。図8は、第2の実施の形態を示す構成図である。カメラ51が、会議に参加している複数の人物を含めた会議室全体を撮影し、その画像（RGB信号またはYUV信号）をフレームメモリ52へ送り、格納する。カメラ51は、広範囲に撮影することができる広角のテレビカメラであるとともに、高解像度の性能を備えたカラーテレビカメラである。ズーム機構や首振り機構を備えてはいず、カメラ51の向きや位置も固定されている。今回フレーム（n）の画像がフレームメモリ52に送られると、フレームメモリ52は今回フレーム（n）の画像を格納するとともに、前回フレーム（n-1）の画像をフレームメモリ53へ送り、そこに格納させる。

【0085】フレーム間差分回路54は、フレームメモリ52から今回フレームを読み出し、また、フレームメモリ53から前回フレームを読み出し、両者の差分を算出する。そして、この差分のフレーム全体にわたった平均値 $e_n$ を算出する。この平均値 $e_n$ を閾値として利用してフレーム間差分の2値化を行い、動画素出力 $M_{x,y}$ を確保する。フレーム内差分回路55は、フレームメモリ52から今回フレームを読み出し、フレーム内差分を算出する。そして、平均値 $e_n$ を閾値として利用してフレーム内差分の2値化を行い、エッジ画素出力 $E_{x,y}$ を確保する。

【0086】重複領域抽出回路56は、動画素出力 $M_{x,y}$ とエッジ画素出力 $E_{x,y}$ との重複部分を算出し、これを動画エッジ画素 $ME_{x,y}$ とする。動画エッジ画素 $ME_{x,y}$ は、会議に参加している複数の人物の実際の存在位置を表している可能性が高い。

【0087】人物抽出回路57は、動画エッジ画素 $ME_{x,y}$ を基にして、人物の存在する可能性の高い領域を特定する。人物判定回路58は、人物抽出回路57で特定された領域の形状が、一般の人物の形状と見なせるか否かを判定する。人物の形状と見なせると判定された場合、顔位置人物位置検出回路59は、人物抽出回路57で特定された領域の形状を基に、顔位置及び身体位置を特定する。

【0088】色情報保存回路62には、前回フレームまでの処理において抽出された人物の顔に関する色パターン情報が格納されている。色パターンマッチング回路60は、顔位置人物位置検出回路59で検出された顔位置における色パターンを、色情報保存回路62に格納されている色パターン情報と比較する。これによって、さらに人物抽出回路57で特定された領域の形状が、人物の形状と見なせるか否かを判定する。登録処理回路61

は、こうした判定の結果に基づき、色情報保存回路62に、今回フレームの処理において抽出された人物の顔に関する色パターン情報を格納する。

【0089】画像切出回路63は、予め設定されている切り出し処理モードに基づき、カメラ51より送られた画像から、画像表示装置に出力されるべき画像範囲を切り出す。画像拡大縮小回路64は、画像切出回路63により切り出された画像を、画像表示装置に合うように拡大または縮小し、一時的に切り出し画像フレームバッファ65に格納する。

【0090】画像圧縮回路66は、切り出し画像フレームバッファ65から画像を読み出し、CCITT-H.261の画像符号化方式により情報量の圧縮を行い、符号化画像を出力する。

$$\epsilon_n = 1/N \sum_{\text{for all pixel}} |P_{n,x,y} - P_{n-1,x,y}| \text{-----} (18)$$

【0094】ただし、Nはフレーム内の画素の総数である。

【S22】フレーム間差分回路54は、画素値 $P_{n,x,y}$ と画素値 $P_{n-1,x,y}$ とを使用して下記式(19)に基づ

$$D_n = |P_{n,x,y} - P_{n-1,x,y}| \text{-----} (19)$$

【0096】【S23】フレーム間差分回路54は、下記式(20)が成立する画素に対しては値1(動画素)を付与し、成立しない画素に対しては値0(静止画素)

$$D_n > TH_{mov} \text{-----} (20)$$

【0098】ここで $TH_{mov}$ は下記式(21)によって決まる値である。

$$TH_{mov} = \alpha_{mov} \cdot \epsilon_n \text{-----} (21)$$

【0100】なお、 $\alpha_{mov}$ は定数であり、これが大きい値であると、動き量の大きい画像を構成する画素に対してだけ値1が付与される。これが小さい値であると、動き量が小さくとも値1が付与されるが、不要なノイズ成分にまで値1が付与されてしまう欠点がある。また、 $TH_{mov}$ を平均値 $\epsilon_n$ を用いて決定することにより、カメラ51で発生する定常的なノイズを除去したり、外部の照明のフリッカによる映像全体の直流分のシフト等に対して安定にフレーム間差分の検出を行い得るようになる。【S24】動画素出力 $M_{x,y}$ において、値1の画素が多数の値0の画素の中に1つまたは2つ孤立して存在するときに、フレーム間差分回路54は、その値1の画素を値0の画素に変換する。また逆に、値0の画素が多数の値1の画素の中に1つまたは2つ孤立して存在するときに、その値0の画素を値1の画素に変換する。これ

$$E_{x,y} < TH_{edg} \text{-----} (22)$$

【0104】ここで $TH_{edg}$ は下記式(23)によって

【0091】こうした構成の詳しい動作を、図9を参照して説明する。図9は、第2の実施の形態における動作を示すフローチャートである。このフローチャートはカメラ51から新たなフレームが入力される毎に実行される。以下、図中のステップ番号に沿って説明する。

【0092】【S21】フレーム間差分回路54は、フレームメモリ52から今回フレーム(n)の画素値 $P_{n,x,y}$ を読み出し、また、フレームメモリ53から前回フレーム(n-1)の画素値 $P_{n-1,x,y}$ を読み出す。ここでx、yはフレーム内の各画素の座標を示す。これらを使用して下記式(18)に基づき、フレーム間差分の平均値 $\epsilon_n$ を算出する。

【0093】

【数18】

き、フレーム間差分値 $D_n$ を算出する。

【0095】

【数19】

を付与し、動画素出力 $M_{x,y}$ として出力する。

【0097】

【数20】

【0099】

【数21】

によって、ノイズが除去される。

【0101】【S25】フレーム内差分回路55は、フレームメモリ52から今回フレーム(n)の画素値 $P_{n,x,y}$ を読み出し、フレーム内差分を算出する。この算出は、多値画像に対するラプラシアン処理方法や方向別のグラディエント抽出処理方法を使用する。そして、平均値 $\epsilon_n$ を閾値として利用してフレーム内差分の2値化を行い、エッジ画素出力 $E_{x,y}$ を確保する。

【0102】すなわち、フレーム内差分回路55は、下記式(22)が成立する画素に対しては値1(エッジ画素)を付与し、成立しない画素に対しては値0(非エッジ画素)を付与し、エッジ画素出力 $E_{x,y}$ として出力する。

【0103】

【数22】

決まる値である。

【0105】

【数23】

$$TH_{edg} = \alpha_{edg} \cdot \epsilon_n \text{ ----- (23)}$$

【0106】なお、 $\alpha_{edg}$  は定数であり、これが大きい値であると、エッジが急峻な画像を構成する画素に対してだけ値1が付与される。これが小さい値であると、エッジが緩やかであっても値1が付与されるが、不要なノイズ成分にまで値1が付与されてしまう欠点がある。また、 $TH_{edg}$  を平均値 $\epsilon_n$  を用いて決定することにより、カメラ51で発生する定常的なノイズを除去したり、外部の照明のフリッカによる映像全体の直流分のシフト等に対して安定にフレーム間差分の検出を行い得る

$$ME_{x,y} = M_{x,y} \cap E_{x,y} \text{ ----- (24)}$$

【0109】図10は、こうして算出された動画エッジ画素 $ME_{x,y}$  の一例を示す図である。図中の升目は画素を示し、空白の升目が値0の画素を表し、斜線部の升目が値1の画素を表す。斜線部の升目が、人物の輪郭線を示している可能性が高い。

【0110】すなわち、本発明とは異なり、フレーム間差分だけにより人物画像を切り取る方式では、フレーム間の時間間隔が長い場合に（例えば数フレーム/秒）、前回フレームの撮影時点から今回フレームの撮影時点までの間に人物が急激に大きく移動したとき、その人物の前回フレームでの画像位置と今回フレームでの画像位置との両方にそれぞれ人物がいると判定されてしまうという問題がある。しかし、本発明のようにフレーム間差分に今回フレームでのフレーム内差分を組み合わせると、そうした2人分の画像位置のうち、前回フレームでの画像位置を消去することが可能となる。

【0111】〔S27〕人物抽出回路57は、動画エッジ画素 $ME_{x,y}$  を基にして、人物の存在する可能性の高い領域を特定する。図11は、ステップS27の詳細な内容を示すフローチャートである。以下、このフローチャートに沿って説明する。

【0112】〔S41〕動画エッジ画素 $ME_{x,y}$  が展開されている画面において、左上の画素を起点にしてラスタースキャンを行う。ラスタースキャンは、画面の各画素を左から右へ、上から下へと順に走査するものである。

【0113】〔S42〕ラスタースキャン対象となった画素が、値1の画素（黒画素）であるか否かを判別する。その結果、値1の画素であればステップS43へ進み、値0の画素であれば、次の画素に対して再度ステップS42を実行する。

【0114】〔S43〕値1の画素であるとされた画素が、画面の右下の最終画素であれば、図11のフローチャートの処理を終了する。以上のステップの実行によって、図10に示した例では、画素p1が検出される。

【0115】〔S44〕値1の画素であるとされた画素が存在する行において、この画素に連結し、かつ、値1

ようになる。

【0107】なおステップS25においても、ステップS24と同様な処理が行われる。〔S26〕重複領域抽出回路56は、動画素出力 $M_{x,y}$  とエッジ画素出力 $E_{x,y}$  とを用いて、下記式（24）に基づき動画エッジ画素 $ME_{x,y}$  を算出する。

【0108】

【数24】

である画素を探す。見つかった場合さらに、見つけられた画素に連結し、かつ、値1である画素を繰り返し探す。こうして同一行において、値1である画素の連続する画素群を検出する。

【0116】図10に示す例では、画素p1と画素p2とからなる画素群が検出される。

〔S45〕ステップS44で得られた画素群を基にして、その直ぐ下の行において、上の行の画素群の幅よりも所定量だけ幅を広くした探索エリアを決定する。この探索エリアは上の行の画素群の直下に位置する。

【0117】図12は、この探索エリアを説明する図である。図中の升目は画素を示し、空白の升目が値0の画素を表し、斜線部の升目が値1の画素を表す。例えば、最上の第1の行L1で10個の画素からなる画素群を検出したとすると、次の第2の行L2では、例えば、左右に3つの画素分だけ増やした16個の画素からなる探索エリアを決定する。第3の行L3でも同様にして探索エリアが決定される。

【0118】〔S46〕ステップS45で決定された探索エリア内に、値1の画素（黒画素）が存在するか否かを判定する。存在すればステップS47へ進み、存在しなければステップS48へ進む。

【0119】〔S47〕ステップS46で存在するとされた複数の画素のうちの最外側の2つの画素で挟まれるエリアを登録し、そのエリア内に位置する値1の画素を値0の画素に変換する。次に実行されるステップS45では、このエリアを「画素群」として扱うようにする。なお、上記エリアが、直ぐ上の行のエリアよりも狭くなる場合には、直ぐ上の行のエリアの幅をこの行のエリアの幅に設定する。

【0120】〔S48〕値1の画素が探索エリア内に存在しない場合、その対象としている行のうち、ステップS44で検出された画素群の下方に位置する部分においては、「白行」と見なされるが、この白行が設定値以上の数にわたって連続して存在するか否かを判定する。存在するならばステップS49へ進み、存在しないならばステップS45へ戻る。

【0121】〔S49〕以上のステップS42からステップS48までの実行によって、人物1人分の画像領域が特定されたことになる。この画像領域を登録する。つぎにステップS42からステップS48までを実行すると、別の人物1人分の画像領域が特定される。そして、全員が特定されると、図11のフローチャートの処理を終了する。

【0122】この結果、図13に示すような登録画像領域が得られる。図13は、図10に示す動画エッジ画素  $ME_{x,y}$  が展開されている画面の例を基に得られた登録画像領域を概ね示している。図13中の升目は数画素分を1目としている。

【0123】図9に戻って、

〔S28〕人物判定回路58は、ステップS49で特定された人物毎の画像領域を基に、その特定された画像領

$$Hn = (x, y, lx, ly, fx, fy, flx, fly) \text{-----} (25)$$

【0127】すなわち、図14に示すように、下辺69と体高70とによって表される人物矩形を作成するとともに、上辺68と顔長さ  $flx$  とによって表される顔矩形を作成する。顔長さ  $flx$  は、上辺68に対する所定比率によって算出される人物の顔の長さである。こうした各矩形において、 $x$ 、 $y$  が、人物矩形の左上角の点の座標を示し、 $lx$  は下辺69の長さを示し、 $ly$  は体高70の長さを示す。 $fx$ 、 $fy$  は、顔矩形の左上角の点の座標を示し、 $flx$  は上辺68の長さを示す。

【0128】図9に戻って、

〔S30〕色パターンマッチング回路60は、顔位置人物位置検出回路59で検出された顔位置における色パターンを、色情報保存回路62に格納されている色パターン情報と比較する。これによって、さらに人物抽出回路57で特定された領域の形状が、人物の形状と見なせるか否かを判定する。登録処理回路61は、こうした判定の結果に基づき、色情報保存回路62に、今回フレームの処理において抽出された人物の顔に関する色パターン情報を格納する。

【0129】図15は、ステップS30の詳しい内容を示すフローチャートである。以下、このフローチャートに沿って説明する。

〔S51〕今回検出された人物矩形が、前回検出された人物矩形とオーバーラップしているか否かを判別する。オーバーラップしていれば、同一人の可能性が高いと判断してステップ52へ進み、オーバーラップしていなければ、別人の可能性が高いと判断してステップS55へ進む。

【0130】〔S52〕今回検出された人物矩形内の顔矩形における色パターンを、オーバーラップしている前回検出された人物矩形の顔矩形における色パターンと比較する。その結果、類似度が大きい場合にはステップS53へ進み、類似度が小さい場合にはステップS54へ進む。

【0131】〔S53〕前回検出された人物矩形内の顔

域が人物に相当することを確認する。

【0124】すなわち、特定された人物毎の画像領域が、例えば図14に示すような形状であったとする。この形状の上辺68の長さと下辺69の長さとの比や、下辺69の長さと体高70の長さとの比を算出し、これらが、前もって設定しておいた値の範囲内に収まっているかどうかを調べる。この結果、設定範囲に収まらない画像領域は捨てる。

【0125】〔S29〕人物と認められる画像領域に関しては、顔位置人物位置検出回路59において、下記式(25)に示すような人物領域登録データ  $Hn$  を人物毎に作成する。

【0126】

【数25】

矩形における色パターンを更新せずに、そのまま、色情報保存回路62に保存する。

〔S54〕色情報保存回路62に保存されている、前回検出された人物矩形内の顔矩形における色パターンに代えて、今回検出された人物矩形内の顔矩形における色パターンを保存する。

【0132】〔S55〕ステップS51～S54の処理を人物矩形毎に行い、検出された全部の人物矩形に対して処理が終了するとステップS56へ進む。

〔S56〕ステップS51においてオーバーラップなしと判別された人物矩形に対して、新たな人物が画面に登場したものとして、新たに、今回検出された人物矩形内の顔矩形における色パターンを、色情報保存回路62に保存する。

【0133】〔S57〕ステップS51においてオーバーラップなしと判別された人物矩形が複数存在する場合、ステップS56の処理を、人物矩形毎に行い、全部の人物矩形に対して処理が終了するとステップS58へ進む。

【0134】〔S58〕ステップS52で類似度大とされた人物矩形に対しては、今回検出された人物矩形内の顔矩形における色パターンを、この顔矩形の位置にある前回検出された人物矩形の色パターンと比較する。その結果、類似度が大きい場合にはステップS60へ進み、類似度が小さい場合にはステップS59へ進む。

【0135】〔S59〕この場合の人物矩形は、ステップ56で登録された人物矩形の残り分であると思われるので、消去する。

〔S60〕この場合、人物が静止状態であるか、または人物以外の静止物を人物と見なしている可能性がある。そこで、その人物矩形に対して、ステップS58によって類似度大と判別された回数を、静止カウンタによりカウントする。

【0136】〔S61〕静止カウンタのカウント値が所

定数 $TH_{stop}$ よりも大きくなった場合には、ステップS62へ進む。

〔S62〕人物がどいた後の壁などの画像を、そのまま人物として扱っていた可能性が高いので、この人物矩形を削除する。

【0137】〔S63〕ステップS52において類似度大と判別された人物矩形が複数存在する場合、ステップS58～S62の処理を、人物矩形毎に行う。人物矩形

$$HR_n = (x, y, lx, ly, fx, fy, flx, fly) \quad (n = 0, \dots, Nm-1)$$

----- (26)

【0140】なお、Nmは、切り出された人物矩形の数である。図9に戻って、

〔S31〕画像切出回路63は人物切り出し矩形 $HR_n$ を受け取り、予め設定されている切り出し処理モードに基づき、カメラ51より送られた画像から、画像表示装置に出力されるべき画像範囲を切り出す。

【0141】図16は、ステップS31の詳しい内容を示すフローチャートである。以下、このフローチャートに沿って説明する。

〔S71〕切り出し矩形パラメータを初期化する。

【0142】〔S72〕人物切り出し矩形 $HR_n$ を読み込む。

〔S73〕切り出し処理モードが「全員包含モード」であるか、「個人切り出しモード」であるかを判別する。「全員包含モード」であるならば、ステップS74へ進み、「個人切り出しモード」であれば、ステップS75へ進む。

【0143】〔S74〕各人の人物切り出し矩形 $HR_n$ を全員分、包含する外接矩形を生成して出力する。

$$\alpha = yL * (1/A) \quad \text{----- (27a)}$$

$$\beta = xL * A \quad \text{----- (27b)}$$

【0148】図16に戻って、

〔S77〕値 $\alpha$ が値 $\beta$ よりも大きいならばステップS78へ進み、値 $\alpha$ が値 $\beta$ 以下であればステップS79へ進む。

【0149】〔S78〕下記式(28a)、(28b)

$$XL = \alpha \quad \text{----- (28a)}$$

$$YL = yL \quad \text{----- (28b)}$$

【0151】〔S79〕下記式(29a)、(29b)に基づき、図17(B)の形状の長さXL、YLを決定する。

$$XL = xL \quad \text{----- (29a)}$$

$$YL = \beta \quad \text{----- (29b)}$$

【0153】なお、切り出し処理モードが「個人切り出しモード」であるときに、ステップS75において、

の全部に対して各処理が終了すると、図15に示すフローチャートの処理を終了する。

【0138】こうした図15に示すフローチャートの処理を実行することにより、登録処理回路61は、下記式(26)で表される最終的な人物切り出し矩形 $HR_n$ を出力する。

【0139】

【数26】

〔S75〕「個人切り出しモード」であり、かつ人物が指定されている場合には、その人物の人物切り出し矩形 $HR_n$ を選択して出力する。

【0144】〔S76〕出力された人物切り出し矩形のアスペクト比は、画像表示装置の画面のアスペクト比と異なっているので、これを調整する必要がある。この調整を、図17を参照して説明する。

【0145】図17(A)は、出力された人物切り出し矩形の形状を表し、図17(B)は、この形状が、画像表示装置の画面のアスペクト比に調整された後の形状を表す。図17(A)の形状の横縦の各長さをxL、yLとし、図17(B)の形状の横縦の各長さをXL、YLとする。画像表示装置の画面のアスペクト比をAとする。

【0146】ここで、下記式(27a)、(27b)に基づき値 $\alpha$ 、 $\beta$ を算出する。

【0147】

【数27】

に基づき、図17(B)の形状の長さXL、YLを決定する。

【0150】

【数28】

【0152】

【数29】

「人物自動選択サブモード」が指定され得るようにしてもよい。この場合には、パラメータを算出比較すること



により特定の人物切り出し矩形HR。を選択し、出力するようにする。パラメータとしては、人物矩形や顔矩形の面積または平均移動量が考えられる。すなわち、これらの値がいずれも大きい場合、それらの値に対応する人物は話し手である確率が高く、こうした人物の画像を切り出すことが有効と思われる。

$$S_n = l_x \times l_y \text{ ----- (30a)}$$

$$FS_n = fl_x \times fl_y \text{ ----- (30b)}$$

【0156】人物矩形の平均移動量 $D_{ave}$ は下記式(31)に基づいて算出する。

$$D_{ave}(n) = D_{body}(n) + \alpha \times D_{ave}(n-1) \text{ ----- (31)}$$

【0158】ただし、 $D_{body}$ は下記式(32)のように表される。

$$D_{body}(n) = \text{abs}(x(n) + l_x(n)/2 - (x(n-1) + l_x(n-1)/2)) \\ + \text{abs}(y(n) + l_y(n)/2 - (y(n-1) + l_y(n-1)/2)) \text{ ----- (32)}$$

【0160】顔矩形の平均移動量 $F_{ave}$ は下記式(33)に基づいて算出する。

$$F_{ave}(n) = F_{body}(n) + \beta \times F_{ave}(n-1) \text{ ----- (33)}$$

【0162】ただし、 $F_{body}$ は下記式(34)のように表される。

$$F_{body}(n) = \text{abs}(f_x(n) + fl_x(n)/2 - (f_x(n-1) + fl_x(n-1)/2)) \\ + \text{abs}(f_y(n) + fl_y(n)/2 - (f_y(n-1) + fl_y(n-1)/2)) \text{ ----- (34)}$$

【0164】かくして画像切出回路63により切り出された画像を、図8に示すように、画像拡大縮小回路64が、画像表示装置に合うように拡大または縮小を行い、一時的に切り出し画像フレームバッファ65に格納する。画像圧縮回路66は、切り出し画像フレームバッファ65から画像を読み出し、CCITT-H.261の画像符号化方式により情報量の圧縮を行い、符号化画像を出力する。

【0165】図18は、カメラ51で撮影された実際の画像から人物矩形が切り出される様子を示す図であり、(A)、(B)、(C)の順に時間経過をしている。

(A)、(B)、(C)の各図において、左端部分はカメラ51の出力画像を示し、中部分はカメラ51の出力画像をそのまま画像表示装置に出力した場合の画像を示し、右端部分はカメラ51の出力画像から人物切り出しを行った結果を画像表示装置に出力した場合の画像を示している。(A)、(B)、(C)の各図における左端部分から分かるように、画面の中で1人の人物が左から右へ、急激に移動している。そうした場合でも第2の実施の形態では確実にその人物像を追尾することが可能である。

【0154】人物矩形の面積 $S_n$ 及び顔矩形の面積 $FS_n$ は下記式(30a)、(30b)に基づいて算出する。

【0155】

【数30】

【0157】

【数31】

【0159】

【数32】

【0161】

【数33】

【0163】

【数34】

【0166】図19は、カメラ51で撮影された複数人の人物画像を含む実際の画像から人物矩形が切り出される様子を示す図であり、(A)はカメラ51の出力画像を示し、(B)は全員を切り出した場合の画像表示装置の画面を示し、(C)、(D)、(E)は特定の1人を切り出した場合の画像表示装置の画面を示す。

【0167】特定の1人を切り出すには、図16のステップ75で示したように、自動的に特定の人物を選択する方法の他に、画像表示装置の画面に対して外部からマウスポインタ等を利用して特定の人物を指定するようにしてもよい。

【0168】

【発明の効果】人物位置検出手段が、カメラから送られた画像を基にして人物の顔位置を検出し、検出した顔位置を基準にして人物の身体位置を検出する。つぎに、画像切出手段が、人物位置の画面上での移動度合いに応じて、人物を含めた表示されるべき画像範囲を決定し、カメラの画像から当該画像範囲を切り出す。すなわち、人物位置の画面上での移動度合いを検出し、この移動度合いが大きいときには、人物を含めた表示されるべき画像範囲を大きく設定し、一方、移動度合いが小さいときに

は、人物を含めた表示されるべき画像範囲を小さく設定する。そして、カメラの画像から、この設定された画像範囲の画像を抽出する。最後に、画面調整手段が、画像切出手段により切り出された画像を画像表示装置の画面の大きさに調整する。

【0169】このようにして、人物抽出を確実にを行い、かつ人物の移動があっても、自動的にズームアウトする方法により見やすい画像表示を可能としている。また、フレーム間差分とフレーム内差分との組み合わせ情報を用いて人物の切り出しを行う。これにより、人物の動きが大きな時でも正確に人物画像を切り出すことができる。

【0170】さらに、人物位置の正確な検出から顔位置の正確な推定も可能となる。これにより、たとえ切り出された人物の動きが小さくなくても、単にフレーム間差分をとって人物の存在領域を検出する従来の方法に比べ、安定した人物追尾を実現できる。

【0171】またさらに、切り出された人物領域の形状が持つ特徴を、一般の人物のそれと比較することにより、人物と人物以外の物との区別を明確に行うことが可能となる。

【0172】また、顔の色パターンを、それまでに記憶していた色パターンと比較することにより、ほとんど静止している人物についても良好に人物切り出しが可能となる。

【図面の簡単な説明】

【図1】本発明の原理説明図である。

【図2】第1の実施の形態の詳しい構成を示すブロック図である。

【図3】顔位置検出回路の動作を説明する図である。

【図4】人物位置検出回路の動作を示す図である。

【図5】切り出し位置決定回路による人物画像切り出し位置の決定手順を示すフローチャートである。

【図6】画面上の人物外接四角形位置を説明する図である。

【図7】送信側から送られた符号化画像を受信側で復号化して画像表示装置に表示した場合の画面の例を示す図である。

【図8】第2の実施の形態を示す構成図である。

【図9】第2の実施の形態における動作を示すフローチャートである。

【図10】動画エッジ画素 $ME_{x,y}$ の一例を示す図である。

【図11】図9におけるステップS27の詳細な内容を示すフローチャートである。

【図12】探索エリアを説明する図である。

【図13】登録画像領域を示す図である。

【図14】特定された人物の画像領域を示す図である。

【図15】図9におけるステップS30の詳しい内容を示すフローチャートである。

【図16】図9におけるステップS31の詳しい内容を示すフローチャートである。

【図17】(A)は、出力された人物切り出し矩形の形状を示す図であり、(B)は、この形状が、画像表示装置の画面のアスペクト比に調整された後の形状を示す図である。

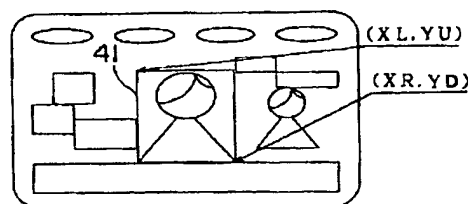
【図18】(A)は、カメラで撮影された実際の画像から人物矩形が切り出される様子を示す第1の図であり、(B)は、カメラで撮影された実際の画像から人物矩形が切り出される様子を示す第2の図であり、(C)は、カメラで撮影された実際の画像から人物矩形が切り出される様子を示す第3の図である。

【図19】(A)は複数の人物画像を撮影したカメラからの出力画像を示す図であり、(B)は全員を切り出した場合の画像表示装置の画面を示す図であり、(C)は特定の1人を切り出した場合の画像表示装置の画面を示す図であり、(D)は他の特定の1人を切り出した場合の画像表示装置の画面を示す図であり、(E)はさらに他の特定の1人を切り出した場合の画像表示装置の画面を示す図である。

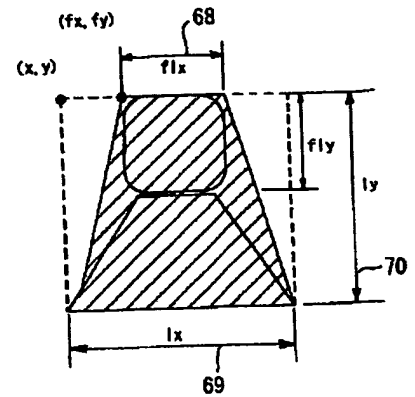
【符号の説明】

- 1 カメラ
- 2 人物位置検出手段
- 3 画像切出手段
- 4 画面調整手段

【図6】



【图14】



```

graph LR
    11[カメラ] --> 12[フレームメモリ]
    12 --> 17[画像切り出し回路]
    12 --> 13[画引きフレームメモリ]
    13 --> 14[顔位置検出回路]
    14 --> 15[人物位置検出回路]
    15 --> 16[切り出し位置決定回路]
    16 --> 18[画像拡大縮小回路]
    18 --> 17
    17 --> 19[切り出し画像フレームバッファ]
    19 -- "送出画像" --> 20[画像圧縮回路]
    20 -- "符合化画像" --> Out[ ]

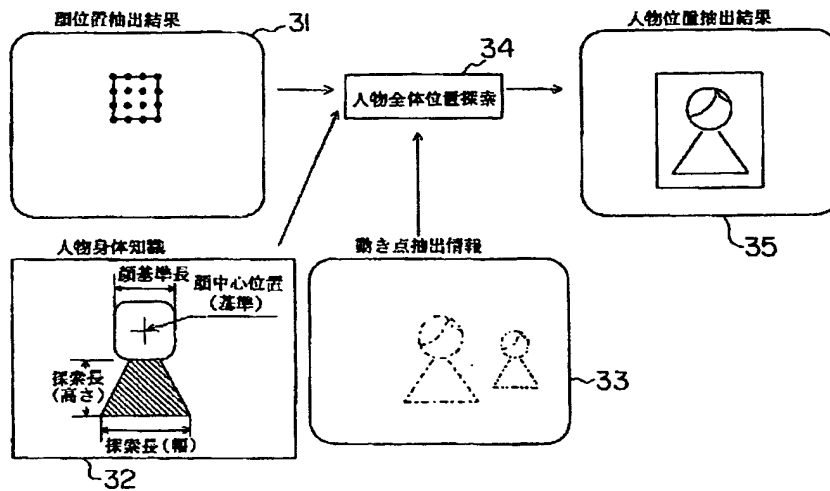
```

```

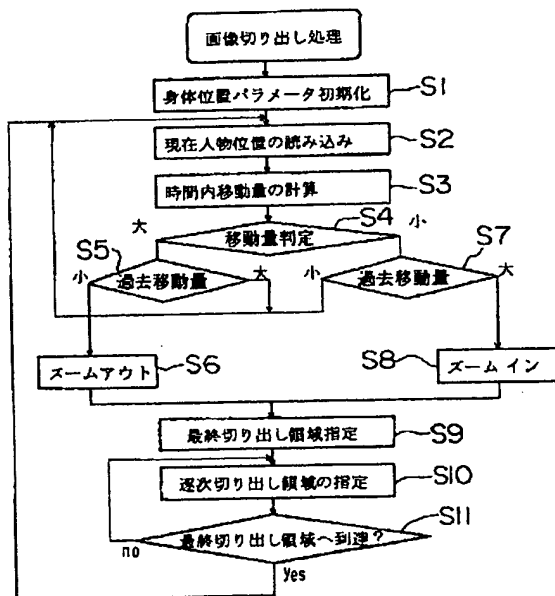
graph LR
    21[21 原画像] --> 23[23 色空間による距離計算]
    22[22 サンプル点配列] --> 25[25 フレーム間差分による動き抽出]
    23 --> 24[24 色抽出点]
    25 --> 26[26 動き抽出点]
    24 --> 27[27 顔位置抽出点群]
    26 --> 27
    27 --> 28[28 抽出した顔位置]
  
```

Figure 1 is a flowchart illustrating the process of face position extraction. The process begins with an original image (21) and a sample point arrangement (22). The original image (21) is processed by color space distance calculation (23) to produce color extraction points (24). The sample point arrangement (22) is processed by frame-to-frame difference motion extraction (25) to produce motion extraction points (26). The color extraction points (24) and motion extraction points (26) are combined to form a group of face position extraction points (27). Finally, the face position is extracted (28).

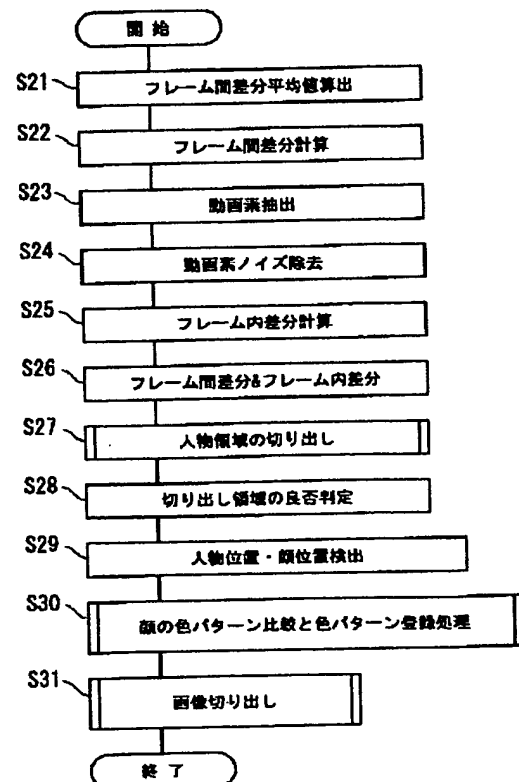
【図4】



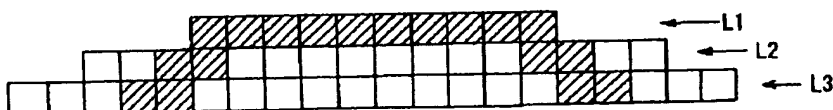
【図5】



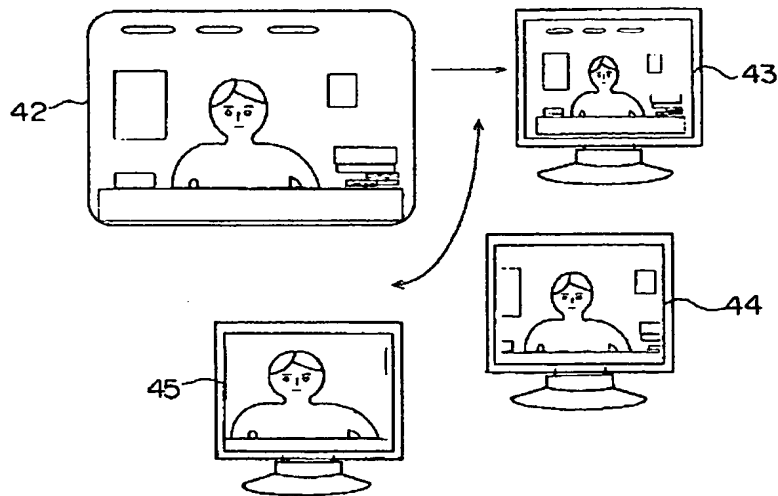
【図9】



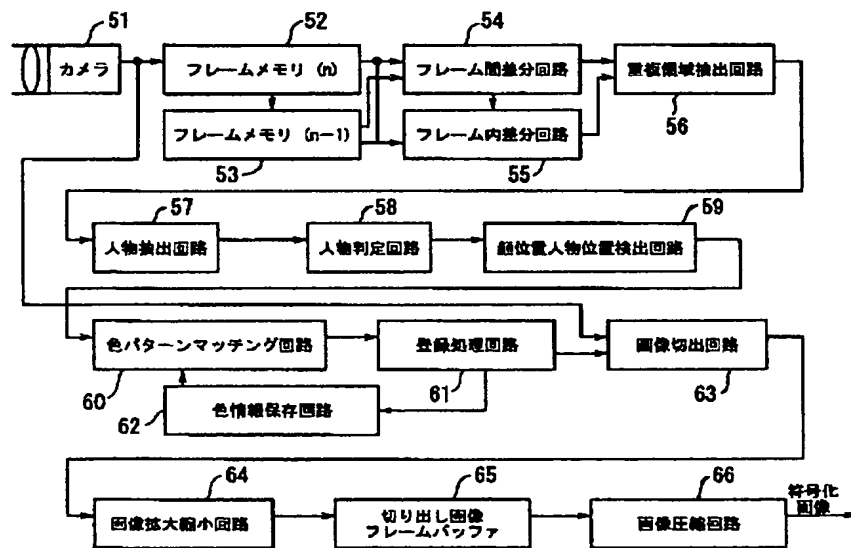
【図12】



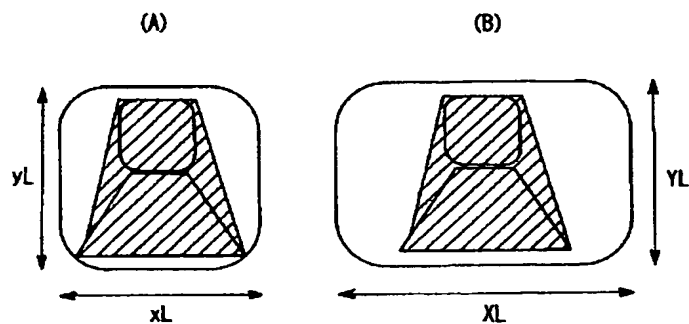
【図7】



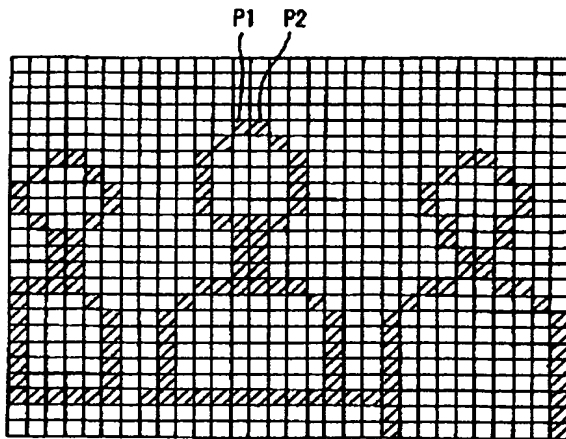
【図8】



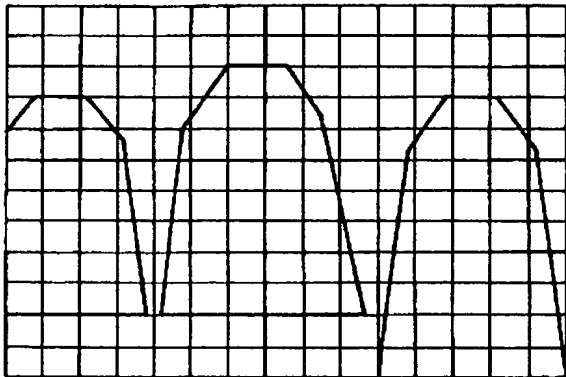
【図17】



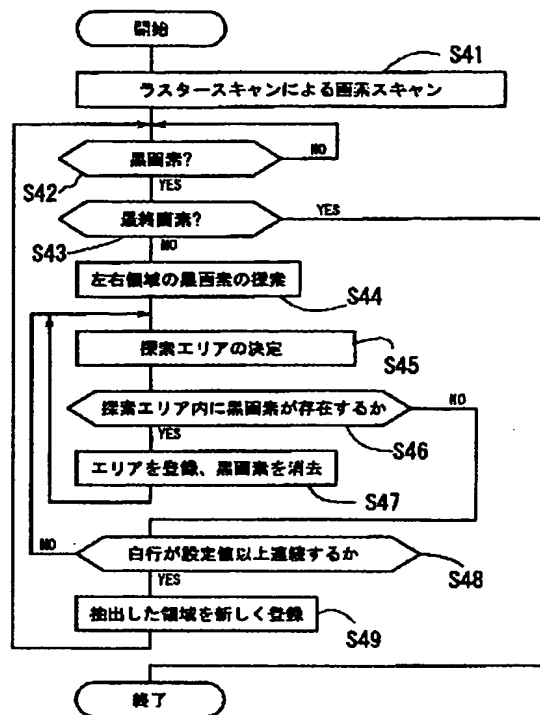
【図10】



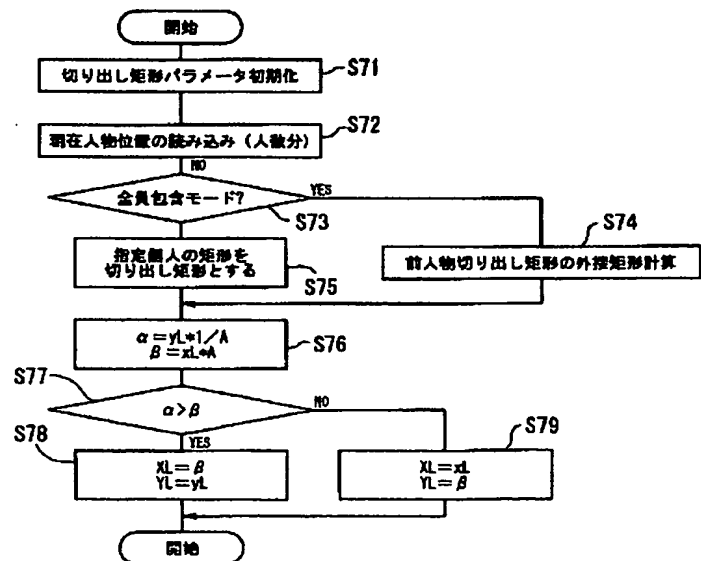
【図13】



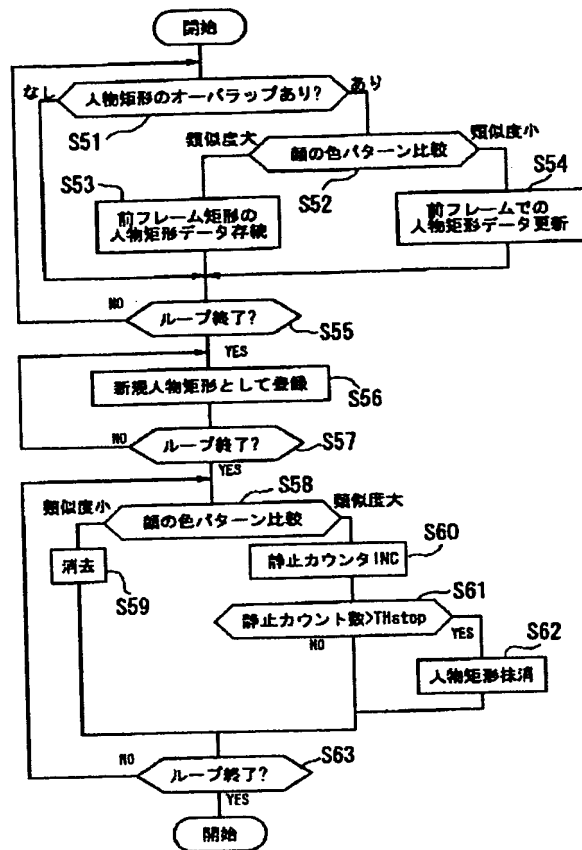
【図11】



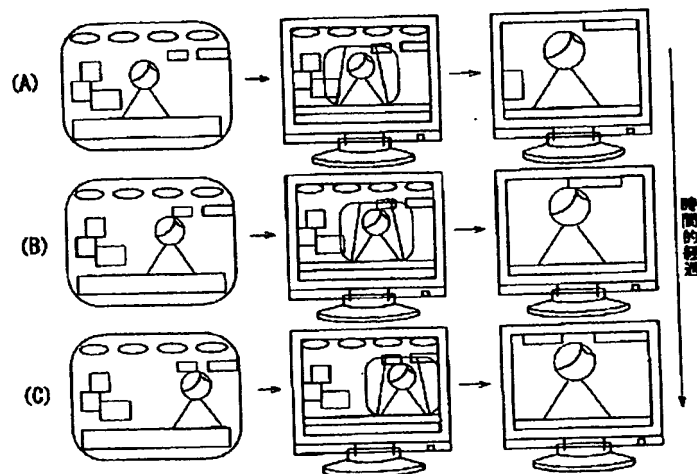
【図16】



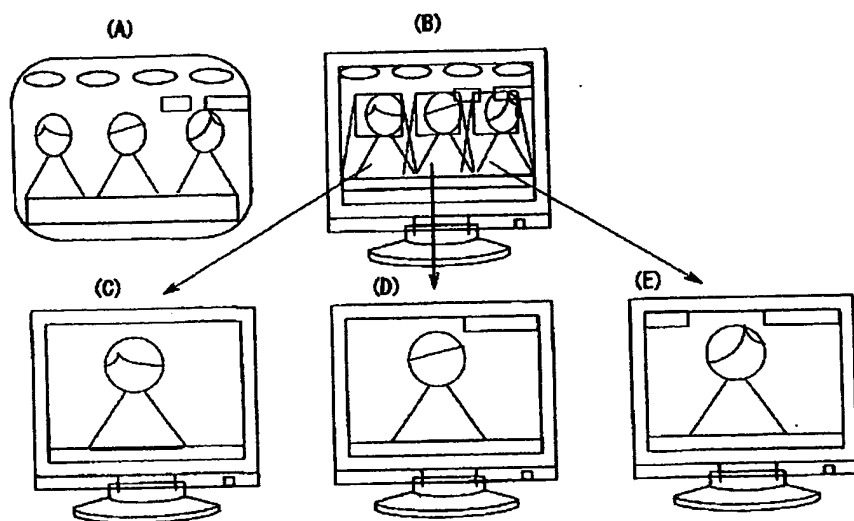
【図15】



【図18】



【図19】





**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☒ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**